

VU Research Portal

Geographic Concentration of Business Services Firms: A Poisson Sorting Mode

Koster, H.R.A.; van Ommeren, J.N.; Rietveld, P.

2011

document version

Early version, also known as pre-print

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Koster, H. R. A., van Ommeren, J. N., & Rietveld, P. (2011). *Geographic Concentration of Business Services Firms: A Poisson Sorting Mode*. (TI Discussion Papers; No. 11-085/3). Tinbergen Institute.
<http://www.tinbergen.nl/ti-publications/discussion-papers.php?paper=1776>

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

TI 2011-085/3
Tinbergen Institute Discussion Paper



Geographic Concentration of Business Services Firms: A Poisson Sorting Model

Hans R.A. Koster
Jos N. van Ommeren
Piet Rietveld

Dept. of Spatial Economics, VU University Amsterdam, and Tinbergen Institute.

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: <http://www.dsf.nl/>

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

Geographic Concentration of Business Services Firms: A Poisson Sorting Model

Hans R.A. KOSTER, VU University Amsterdam*

Jos N. VAN OMMEREN, VU University Amsterdam

Piet RIETVELD, VU University Amsterdam

This version: May 17, 2011

Abstract

This paper examines the effects of specialisation (within-sector clustering) and diversity (between-sector clustering) on business services profitability and location choice. We apply a semiparametric Poisson sorting model allowing for firm-specific effects. We find that for most firms, profitability of business services firms is substantially higher close to specialised clusters of business services firms. A standard deviation increase in business services specialisation of a location leads to on average a 40 percent increase in the probability that a business services firm locates there, supporting theories of Marshall, Arrow and Romer. It is also profitable for most business services firms to locate near a group of firms that belong to the same sector, not necessarily business services firms, so diversity is negatively related to location decisions. Almost all firms either benefit from within-sector clustering or between-sector clustering. Within-sector clusters are particularly profitable for large mature firms, whereas between-sector clusters are relatively more profitable for smaller innovative firms.

Keywords: Sorting; Agglomeration Economies; Specialisation; Diversity; Heterogeneity; Semiparametric Estimation

* Corresponding author. Department of Spatial Economics, VU University, De Boelelaan 1105 1081 HV Amsterdam, e-mail: h.koster@vu.nl. We would like to thank NICIS-KEI for financial support.

1. Introduction

Concentration of firms at certain places is too large to be explained by exogenous differences in natural advantages only (Ellison et al., 2010). It is very likely that firms tend to concentrate because of local interactions with other firms. Proximity to others reduces transport costs of goods, people and ideas (Marshall, 1920; Ellison et al., 2010). Marshall (1920), Arrow (1962) and Romer (1986) (MAR) argue that *specialisation* (or within-sector clustering) generates knowledge spillovers between firms. Innovators' ideas will be copied, imitated and improved by their neighbours. It is argued that in particular nearby firms of the same type will be able to copy these technologies (Porter, 1990; Audretsch and Feldman, 1996). Besides these dynamic externalities, static externalities also play a role. For example, a cluster of a specialised industry induces a market for specialised suppliers and generates a thick labour market. In contrast, Jacobs (1969) argues that a *diverse* economy will attract firms because it may lead to a more diverse portfolio of customers and exchange of tacit knowledge, as knowledge and technologies from different sectors may be recombined into new combinations (see Schumpeter, 1942; Fujita et al., 1999). This may be labelled diversity or between-sector clustering.

An extensive literature, started by the seminal work of Glaeser et al. (1992), reviews the effects of specialisation and diversity on growth of cities, regions and even countries (see for a meta-analysis De Groot et al., 2009). In the current literature, the emphasis is on the effects of these forces on *new* firms (see Head et al., 1995; Guimaraes et al., 2000; 2002; Rosenthal and Strange, 2003; Devereux et al., 2007; Arzaghi and Henderson, 2008). Although a focus on new firms is useful as it simplifies the econometric procedure, it also means that we know little to what extent these forces contribute to the profitability and therefore location choices of *established* firms (an important exception is Ellison et al., 2010). The difference between new and established firms is fundamental, not only because new and established firms differ strongly in terms of observable characteristics (e.g. new firms are relatively small), but also because new firms systematically choose different locations than established firms (Henderson et al., 1995; Duranton and Puga, 2001). In the current paper we focus on local interactions by business services firms. So, we aim to examine how the

profitability of established business services firms depends on specialised business services clusters (e.g. the local employment share of business services firms) and on between-sector clusters (diversity), so the extent to which the local composition of industrial sectors mirrors the composition of the broader economy.¹

In the current study we pay much attention to firm heterogeneity. It may be expected that the contribution of specialisation and particularly diversity to profits differs strongly even between firms belonging to the same business services sector. This is in line with Baldwin and Okubo (2006) who show that in a theoretical model with heterogeneous firms, a regional policy induces more productive firms to move to the city centre, whereas less productive firms locate in the periphery. Duranton and Puga (2001) argue that small new firms have to experiment with production processes and therefore locate at diversified locations. A more mature firm with an optimised production process will then relocate to a specialised location with firms that have the same production process.² So, it is less informative to provide the average effect of diversity, also because it is plausible that many firms negatively rather than positively value diversity.³ Given heterogeneous effects, it is likely that different types of firms (e.g. large or small, different subsectors) sort themselves into different locations.

We estimate a heterogeneous firm sorting framework, in which we assume profit maximising firms, to measure the firm-specific effects of specialisation and diversity. The essential assumption of the proposed framework is that in equilibrium, firm-specific profit advantages related to location are continuous over space. This leads to a semiparametric conditional Logit model (CLM) that one may estimate by means of a

¹ Business services are defined as professional services, such as consultancy, financial services, but also include new-technology based services such as telecommunications and Research and Development (Muller and Zenker, 2001). Current empirical studies that pay specific attention to the impact of diversity and specialisation on location decisions are on an aggregate spatial level and focus mostly on extra-metropolitan agglomeration economies in the manufacturing sector, whereas it is suggested that, especially for business services firms, local interactions are more relevant (Muller and Zenker, 2001; Duranton and Overman, 2005; Drennan and Kelly, 2010; Ellison et al., 2010). For example, Arzaghi and Henderson (2008) show that interactions between advertising agencies mainly take place within a couple of hundred meters.

² Possibly because of firm heterogeneity, different studies arrive at opposite conclusions about whether specialisation and diversity attract or repel firms because there is no distinction made between established and new firms. For example, most *established* communication agencies are specialised in the communication for certain sectors (government, healthcare, pharmaceutical), but *new* communication agencies usually do business for many sectors.

³ By focusing on firm-specific effects combined with an instrumental variables approach, we also avoid the reflection problem (e.g. Manski, 2000). This problem occurs in linear models with homogeneous firms. Then, one cannot distinguish between individual and mean behaviour.

locally weighted Poisson regression. This approach reveals firm-specific effects for specialisation and diversity without making assumptions regarding the distribution of these effects.⁴ Our estimation procedure consists of two stages and can be carried out using standard statistical software packages.⁵

In the first stage of our empirical application, we estimate firm-specific effects of specialisation and diversity on location choices, so we aim to disentangle the effect of these potentially endogenous measures from (unobserved) natural advantages. We correct for endogenous measures using a control function approach that relies on instruments external to the model.⁶ The results are compared to an approach employing internal instruments, which rely on within-region variation (see Bayer and Timmins, 2007). In the second stage, we estimate the effect of firm characteristics, and in particular workforce size on the firm-specific effects estimated in the first stage, which helps us to explain why large and more mature firms choose specialised locations, whereas small innovative firms are more profitable in diversified locations.

This paper proceeds as follows. Section 2 outlines our equilibrium sorting model for firms, followed by the estimation procedure in Section 3. In this section we also show when the equivalence relation between locally weighted CLM and Poisson holds. Section 4 discusses our dataset and the instruments used. In Section 5 we discuss the empirical results. Standard parametric techniques are compared with the results of our semiparametric approach. We pay especially attention to the heterogeneity in the firm-specific effects. Section 6 concludes.

⁴ It has been shown that there is an equivalence relation between a standard CLM and Poisson regression (Guimarães et al., 2003; Schmidheiny and Brülhart, 2011). We show that the equivalence relation also holds in a semiparametric setting, but maybe surprisingly, only when assumptions are made on the values of the weights. It appears that these assumptions imply a '*k*-nearest neighbour approach', frequently applied in the semiparametric literature, which has an intuitive economic interpretation.

⁵ We will not use the methodology of Ellison et al. (2010) because their methodology requires data that are only available for manufacturing firms (e.g. patent data), while we focus on business services firms. A large advantage of our approach is that it identifies firm-specific profit parameters, whereas the methodology of Ellison et al. (2010) is not directly related to the profit function of specific firms.

⁶ In many studies there is not a proper distinction made between local interactions and (unobserved) natural advantages, which generally leads to overstated effects of agglomeration. Ellison and Glaeser (1999) conjecture that at least half of the observed concentration is due to natural advantages.

2. A heterogeneous sorting model for firms

In our equilibrium framework, we assume that firm i chooses a small location j that maximises profit π_j^i , which is a function of locational attributes \mathbf{X}_j (e.g. employment density) as well as a specialisation measure ρ_j and a diversity index σ_j , where $i = 1, \dots, I$ and $j = 1, \dots, J$. Then:

$$(1) \quad \max \pi_j^i = \alpha_0^i + \mathbf{X}_j' \boldsymbol{\alpha}^i + \beta^i \rho_j + \gamma^i \sigma_j + \epsilon_j^i,$$

where ϵ_j^i denotes an idiosyncratic profit advantage (or disadvantage) of firm i at location j . The α 's, β 's and γ 's are firm-specific coefficients to be estimated, which depend on the specific underlying technology of the firm. So, the profit parameters are some function of underlying firm characteristics:

$$(2) \quad \{\boldsymbol{\alpha}^i, \beta^i, \gamma^i\} = \Theta(\mathbf{Z}^i, \xi^i),$$

where $\Theta(\cdot)$ is some function of observable characteristics \mathbf{Z}^i , such as workforce size, and an unobservable characteristic ξ^i . In (1) we assume an additive profit function but it is otherwise very general, as the coefficients are firm-specific.

To estimate sorting models based on (1), one needs to impose some restrictions. We assume that in equilibrium, so after sorting, coefficients are continuous over space. More formally, we assume that for all coefficients α and β the absolute difference between coefficients of two firms, so $|\alpha^i - \alpha^{\tilde{i}}|$ and $|\beta^i - \beta^{\tilde{i}}|$, are a positive function of $d_{j\tilde{j}}$, where $d_{j\tilde{j}}$ denotes the distance between the locations of firms i and \tilde{i} . We think this is a natural assumption, as firms with similar coefficients will sort themselves across space. It is quite common to make a similar assumption, for example in the econometric literature on spatial discontinuity designs (e.g. Black, 1999; Bayer et al., 2007). Sorting of households with similar preferences and sorting of firms with similar technologies is also an outcome of many theoretical urban equilibrium models (e.g. LeRoy and Sonstelie, 1983; Glaeser et al., 2008). In the special case where $d_{j\tilde{j}} = 0$, we assume $\boldsymbol{\alpha}^i = \boldsymbol{\alpha}^{\tilde{i}}$ and $\beta^i = \beta^{\tilde{i}}$. So, firms that locate at exactly the same location (so, for example within the same building) are assumed to have identical coefficients. Note that the latter assumption is hardly restrictive because locations are small,

so there are few firms per location.⁷ We are not too concerned when the latter assumption does not strictly hold: when firms with different coefficients locate at j , the average coefficient for a location is obtained, and therefore the firm-specific parameters of the profit function are *approximately* identified. When we assume that ϵ_j^i has an Extreme Value Type I distribution, it can be shown that:

$$(3) \quad \Pi_j^i = \frac{e^{\alpha_0^i + X_j^i \alpha^i + \beta^i \rho_j + \gamma^i \sigma_j}}{\sum_{\tilde{j}} e^{\alpha_0^i + X_{\tilde{j}}^i \alpha^i + \beta^i \rho_{\tilde{j}} + \gamma^i \sigma_{\tilde{j}}}},$$

where Π_j^i denotes the probability that firm i chooses location j , and \tilde{j} are all locations including j . So, (3) implies a firm-specific conditional Logit model. As is well known, this model is only identified given a normalisation assumption, so the coefficients are identified up to a constant and must be interpreted as the effects on profitability, rather than on absolute levels of profits.

3. Estimation procedure

3.1 The first stage: a local Poisson approach

Given (3), firm sorting models can in principle be estimated using a CLM. However, in practice there are some serious difficulties when employing this estimation procedure. The main disadvantage is the very computational intensive procedures to arrive at results. Firms have many locations to choose from and therefore the choice set grows almost infinitively large when locations are defined more narrowly.⁸ Guimarães et al. (2003; 2004) and Schmidheiny and Brülhart (2011) address this issue by noting that estimation of a CLM and a Poisson model is equivalent given the assumption that firms are homogeneous.⁹

To estimate Poisson models is computationally less intensive and therefore often the preferred method (see

⁷ In our data, 68 percent of business services firms reside in locations where no other business services firms are located. 16 percent are located in areas where only one other business services firm is located and only 16 percent reside at locations where two or more other business services firms are located.

⁸ One solution is to focus on large aggregated regions, but then local factors that determine locational choice are not identified. McFadden (1978) argues that the CLM could still be applied given a large number of regions, by randomly selecting a choice subset (e.g. locations) which is obtained from the full choice set. This idea is much applied in locational choice models, including recent studies about sorting of households (see, among others, Bayer et al. 2007, Ioannides and Zabel, 2008). Still, to estimate the conditional Logit model is quite computationally intensive, even if a number of restrictive assumptions concerning the profit/utility function are made.

⁹ Schmidheiny and Brülhart (2011) show that, although coefficients are the same, predictions of the CLM and the Poisson model may differ substantially if the number of locations is limited, as elasticities are then different for both models. In the current study, the number of locations is quite substantial, so this issue does not apply here.

for applications, Gabe and Bell, 2004; Holl, 2004; Woodward et al., 2006; Arzaghi and Henderson, 2008). We aim to relax this assumption. At first sight, it may seem impossible to estimate a heterogeneous sorting model by means of a Poisson model, as the homogeneity assumption implies aggregating of individual firm data (Defever, 2006). However, it can be shown that one may still study heterogeneity in firms behaviour using a locally weighted Poisson-approach, given certain restrictions on the weights. We then apply an insight of the hedonic house price literature, where individual household characteristics are used in a second stage to derive the household-specific willingness to pay for housing attributes (see Ekeland et al., 2004; Bajari and Benkard, 2005; Bajari and Kahn, 2005).

In the current paper, we allow for firm heterogeneity by estimating locally weighted Poisson models (in the first stage). ‘Local’ implies that for each location a weighted Poisson model is estimated. In the empirical literature, several ways to determine the local weights are employed. We employ a k -nearest neighbour approach. This approach estimates for each location j a Poisson regression where the weight equals one when location j is part of the subset of k nearest locations and zero otherwise (Fotheringham et al., 2002). The choice of k , also known as the window size, is equal for all locations.¹⁰ In Appendix A we show that the locally weighted Poisson model and the locally weighted CLM are equivalent, but only when weights are either zero or a constant (normalised to one), consistent with the k -nearest neighbour approach. So, our approach implies that firms only take a subset of k nearest locations into account. This assumption seems realistic, especially for established business services firms as the large majority of these firms (about 75 percent) when they relocate, do so within a couple of kilometres from the original location (Van Dijk and Pellenbarg, 2000).

¹⁰ In principle, one may allow for window sizes that vary between locations (for example, using a fixed distance). However, this will lead to excessive smoothing in areas where there are many observations and to a high variance in areas with sparse data (see Stute, 1984; Cho et al., 2009). McMillen and Redfearn (2010) therefore argue that a k -nearest neighbour approach is preferred. Another issue is that the weighting scheme may seem restrictive, as we have to use discrete weights instead of weights that are continuous over space. However, it is often argued that this will not influence the results that much (e.g. McMillen, 2010). The choice of window size k is much more important. It is also noted that given a large number of locations, our estimation procedure will lead then to approximately continuous coefficients over space, in line with Section 2, where we assumed that profit coefficients are continuous over space.

3.2 Addressing endogeneity in the first stage

It may be argued that the specialisation measure ρ_j is endogenous with respect to location j , as unobserved natural advantages may cause clustering of business services firms. For example, favourable municipality-specific tax regimes may attract business services to certain locations. To correct for endogeneity, we employ a control function approach (see Blundell and Powell, 2003; Yatchew, 2003; Blundell and Powell, 2004).¹¹ This approach treats endogeneity as an omitted variable problem, comparable to Heckman's correction for selectivity bias, through the introduction of an appropriately estimated control function (Heckman, 1979). An important restrictive assumption of the control function approach is that the endogenous variables should be continuously distributed, which is fulfilled in our application. The control function approach is preferred to two main alternative approaches to correct for endogeneity such as IV and plugging in fitted values (Blundell and Powell, 2003).

The procedure to apply the control function is to first regress the endogenous independent variables on all exogenous independent variables and instruments (using the nearest neighbour approach outlined in the previous section). In the second step one regresses the number of firms per location on all independent variables and the predicted first step errors (again using the nearest neighbour approach).

3.3 The second stage

We are particularly interested the effect of firm characteristics on the profitability of specialisation and diversity, to test whether larger firms choose specialised locations, whereas small firms prefer diversified locations. We therefore regress the estimated profit parameters of specialisation and diversity on firm characteristics workforce size and (sub)sector using ordinary least squares. So, it is assumed that $\{\hat{\beta}^i, \hat{\gamma}^i\} = \mathbf{Z}^i \theta + \xi^i$, where θ are the coefficients to be estimated and the unobserved characteristic ξ^i enters additively, similar to Bajari and Kahn (2005).

¹¹ For recent applications of the control function approach, see Guevara and Ben-Akiva (2006) and Petrin and Train (2010).

4. Data and instruments

4.1 Dataset and regional context

Our dataset contains information about characteristics of *all* establishments in 2005 in the NUTS3-region Zuid-Holland, located in the west of the Netherlands.¹² This region is 2,860 square kilometres and has the highest population density in the Netherlands (about 1,200 people per square kilometre). It includes Rotterdam, the second largest city of the Netherlands, The Hague, where the national government is located, but also cities such as Leiden and Delft. Zuid-Holland covers about 20 percent of national economic activity. The data on firm locations comes from administrative sources and is very reliable, as Dutch firms are obliged by law to provide this information. We have information on the firm's exact location on a 6-digit postcode level (PC6) and the number of employees, as well as detailed information on sector. A PC6 area is small (comparable to the size of a census block in the United States). It includes on average 17 workers of which 3 are employed by a business services firm. About 15 percent of the firms are business services firms and about 17 percent of the total workforce is employed in business services firms (see Table 1).

Table 1: Firms and Employment in Zuid-Holland

	Firms		Employment	
	Number	% of Total	Number	% of Total
Total	62,424	100.00	1,477,354	100.00
Rotterdam	11244	18.01	322,688	21.84
The Hague	7583	12.15	232,630	15.75
Leiden	2110	3.38	53,874	3.65
Dordrecht	1934	3.10	53,990	3.65
Delft	1660	2.66	49,422	3.35
Zoetermeer	1658	2.66	48,094	3.26
Business Services	9,172	14.69	249,722	16.90
Rotterdam	1,994	17.73	64,831	20.09
The Hague	1,456	19.20	54,335	23.36
Leiden	294	13.93	5,440	10.10
Dordrecht	280	14.48	6,452	11.95
Delft	342	20.60	10,587	21.42
Zoetermeer	350	21.11	11,306	23.51

NOTE: Firms refer to establishments with at least 3 employees.

¹² In the remainder of this paper, we label establishments as firms.

One issue to consider here are the large number of small firms. About 50 percent of the observations are firms with one or two employees. In administrative sources, small organisations are frequently fiscal entities that do not require the physical presence of employees. Furthermore, a self-employed entrepreneur can work from home. Then, the choice where to live and work is a joint-decision. We therefore only select business services firms with three or more employees, 9,170 observations in total.¹³ Conditional on this selection, the average firm size is 25 employees.

Following Duranton and Overman (2005), we control for zoning (which restricts the choice set of locations) by allowing firms only to choose from locations where at least one firm is located (which does not have to be a business services firm). We then have 21,465 locations, so the average number of business services firms per location is 0.427. 76 percent of the locations do not host any business services firm, 16 percent only one business services firm and only 7.7 percent two business services or more. The average distance of a business services firm to all other business services firms is 21.22 kilometres in the Zuid-Holland region.

4.2 Defining the explanatory variables

We aim to measure the effects of specialisation and diversity on profitability and therefore on location choice. We will use measures which are continuous over space, to overcome biases inherent to the use of indices based on discrete spatial units (Duranton and Overman, 2005; Ellison et al., 2010). Our starting point is to measure the weighted average of the number of jobs in sector G located in the neighbourhood of location j using an exponential distance decay function (see Fujita and Ogawa, 1982; Lucas, 2001; Lucas and Rossi-Hansberg, 2002). Formally:

$$(4) \quad \omega_{Gj} = \delta \sum_j^J e^{-\delta d_{jj}} \ell_{Gj},$$

¹³ We use the Standard Industrial Classification (SIC) to select business services from the complete dataset of firms.

where ω_{Gj} denotes the weighted employment of location j of sector G , $\ell_{G\tilde{j}}$ is the number of jobs in sector G in location \tilde{j} and δ denotes a fixed decay parameter. We note that distance $d_{j\tilde{j}}$ is measured in kilometres. Specialisation is a measure of concentration of an industry in a city. Our specialisation measure is defined as:

$$(5) \quad \rho_{Gj} = \frac{\omega_{Gj}}{\sum_{\tilde{G}} \omega_{\tilde{G}j}},$$

which is the share of weighted employment in sector G over the total weighted employment in location j . In our application, we investigate the tendency of business services to cluster. The business services specialisation measure ρ_j therefore refers to the local share of business services employment (so in our analysis $\rho_j = \rho_{Gj}$, where $G = \text{business services}$).

To measure between-sector diversity σ_j we use a (continuous) measure, based on the diversity index of Duranton and Puga (2000):

$$(6) \quad \sigma_j = 1 / \sum_{\tilde{G}} |\rho_{\tilde{G}j} - \rho_{\tilde{G}}|,$$

where $\rho_{\tilde{G}} = \sum_j \ell_{\tilde{G}j} / \sum_j \sum_{\tilde{G}} \ell_{\tilde{G}j}$. This index corrects for differences in sectoral employment shares at the regional level. The coefficient related to the diversity index will, conditional on our specialisation measure, represent the profit advantage of a firm to have a diverse portfolio of firms in the neighbourhood. When it is negative, business services will benefit more from one type of firm in the neighbourhood (e.g. only manufacturing or only government). Specialisation and diversity are not measures that are orthogonal to each other (see Figure B1).¹⁴ To construct the specialisation and diversity indices, we assume that $\delta = 2.5$, so that on average more than 90 percent of the weight is within 1.5 kilometres of location j .¹⁵ To facilitate interpretation, we transform ρ_j and σ_j to have zero mean and unit standard deviation. Figures 1 and 2 present the geographical distributions of these two (standardised) explanatory variables.

¹⁴ In particular, we note that when the specialisation measure ρ_{Gj} approaches one (so $\rho_{\tilde{G}j} = 0$), σ_j approaches $1 / (2 - 2\rho_G)$. So, in this specific case there is a deterministic relation between specialisation and diversity (which is intuitive). However, there are few regions with very high shares of business services, so both variables refer to different concepts and the effect of these variables can be separately identified.

¹⁵ Recent empirical research provides evidence that interactions take place very locally, within 2 kilometres of the location of the firm. See, among others, Rosenthal and Strange (2003; 2008), Arzaghi and Henderson (2008).

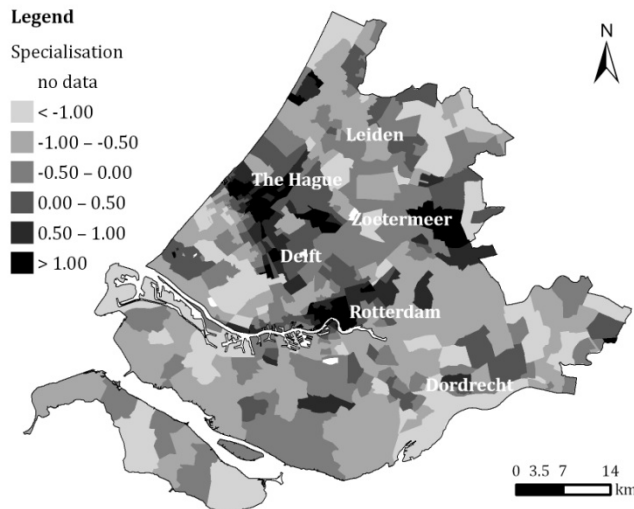


Figure 1: Pattern of specialisation

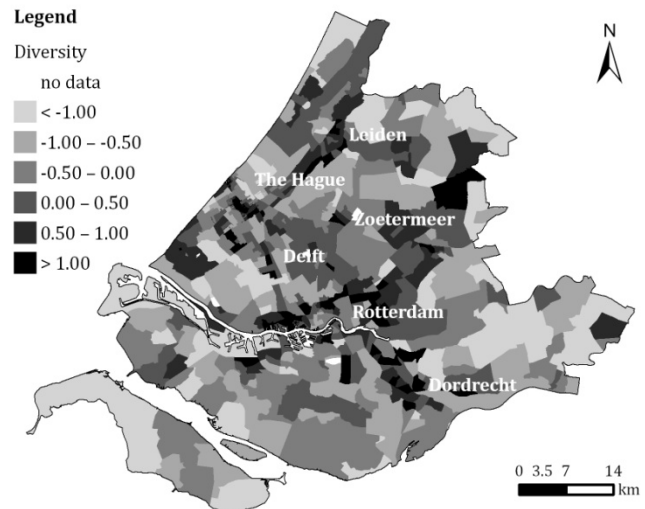


Figure 2: Pattern of diversity

Business services firms are concentrated in the corridor Rotterdam–Delft–The Hague (see Figure 1). In these cities more than 20 percent of the total employment is in the business services sector, which is about 8 percentage points higher than for regions outside this corridor.¹⁶ Rural areas in the south, the east and the north experience relatively low concentration of business services (see Figure 1). These areas are mostly specialised in rural activities, along with regular retail activities. The pattern of diversity is somewhat similar to the specialisation pattern, although there are some differences (see Figure 2). Generally, urban areas are more diversified than rural areas.¹⁷

In the analysis, we control for other locational attributes. In particular, we condition on employment density, so we include the number of employees in the neighbourhood of the firm (which is endogenous), using the distance decay function in (4). This variable may be interpreted as a density variable that captures the combined positive Marshallian externalities and negative effects of higher rents in high density areas,

¹⁶ Rotterdam hosts a large cluster of maritime business services which provide services to port-related manufacturing firms, such as the petrochemical industry. The concentration of business services near Delft may be explained by the presence of a technical university, as Research and Development firms may find it attractive to locate near a university (Jaffe, 1989; Woodward et al., 2006).

¹⁷ Rotterdam, the largest city in the region, has a diversified portfolio of employment, maybe because of the mix of port-related industries, urban amenities and business services. The Hague is only partly diversified, because some areas are specialised in public services. Medium-sized cities such as Dordrecht and Gouda are generally diversified.

congestion and other crowding effects (Bartik, 1985; Guimarães et al., 2000; Barrios et al., 2006).¹⁸ We also control for the distance to the nearest highway ramp, station, hectare of water and open space. To include the latter is relevant because of zoning regulations, so firms are not allowed to locate everywhere (for example, close to residential areas, or protected agricultural areas). We correct for the size of the postcode area, because larger regions are likely to host more firms.¹⁹ The descriptives of those explanatory variables are presented in Appendix B. To control for unobserved heterogeneity at a local level, we also include a polynomial function of geographic coordinates. This is a flexible way to control for unobserved spatial factors, and in line with the assumptions that, in equilibrium, profit coefficients are continuous (see Cheshire and Sheppard, 1995).²⁰

4.3 Instruments

Our business services specialisation measure and the control variable employment density are potentially endogenous. We have to seek instruments that are correlated with the pattern of clustering of (business services) firms, but uncorrelated with unobserved attributes that directly affect the location choice (and profits) of business services firms.²¹ First, we include a dummy indicating whether the location was bombed in 1940. In the beginning of the Second World War, the complete city centre of Rotterdam was destroyed. Afterwards, the city centre was completely redeveloped into the most dense city centre in the Netherlands, hosting many business services. As a second instrument we measure whether a location was below water in

¹⁸ Multicollinearity is not to be a problem here: the correlation between specialisation and diversity is only 0.18 and between specialisation and employment density it is 0.56. The correlation between diversity and employment density is also quite low and equal to 0.23.

¹⁹ Area size is potentially endogenous, as areas that attract many firms may be expanded. However, excluding this variable will lead to very similar results. The area size may differ substantially between urban and rural areas. As an illustration: areas that do not host any business services firms are on average 9.34 hectare, whereas areas that host business services firms are on average 4.25 hectare.

²⁰ It is noted that one may only include variables that have sufficient *local* variation to avoid local perfect multicollinearity (Wheeler and Tiefelsdorf 2005; McMillen and Redfearn, 2010). We therefore include a second-order polynomial of the geographic coordinates. In principle, it is also possible to use higher-order polynomial, but this leads to severe local multicollinearity.

²¹ Bayer and Timmins (2007) argue that instruments for agglomeration measures arise naturally out of the sorting model. Therefore, there is no need for external instruments. However, the validity of such internal instruments depends on the variation in the effects of specialisation and diversity. If there is not enough variation, the estimates are not robust to misspecifications of the profit function. We will compare the internal instrument approach with our approach in Section 5.4.

1830, so whether the land has been reclaimed. For these areas, employment density was strictly zero in 1830. So, due to autocorrelation in employment density, we may expect that specialisation and employment density are lower in these areas. We also include an instrument that captures the distance to the nearest station in 1850, while, and this is important, controlling for the current distance to the nearest station. Railway stations were an important factor enforcing agglomeration of firms and people around that time (Ciccone and Hall, 1996). As a fourth instrument, we use population density of municipalities *in 1830*.²² The instrument's validity rests on the assumption that population density in 1830 is unrelated to current firms' location decisions (and therefore firms' profit), but has a causal effect on the current agglomeration *pattern* (see also Ciccone and Hall, 1996; Rice et al., 2006; Combes et al., 2008). This instrument is strong as population density is strongly autocorrelated and (current) population and employment densities are positively correlated (McMillen and McDonald, 1998).²³ The population density instrument (but not the other three) may be criticised using the argument that there are serially correlated unobservable variables which invalidate the instrument. In the case of business services firms, this argument seems not particularly convincing as this type of firm did not exist in 1830 and does not have households as customers (furthermore, we control for geographic coordinates which captures any positive effect of being close to the residence location of its workforce).

Thus, we arrive at four instruments, which exceeds the number of endogenous variables. This is useful although the first two instruments only have local impact. For example, the instrument whether an area was bombed only applies for locations near Rotterdam. Still, as these two instruments locally have a very strong impact, they contribute to the identification of the firm-specific effects of specialisation and employment density.

²² Municipalities in 1830 were much smaller and do not overlap with the current ones. Zuid-Holland, the region which our data refer to, consisted in 1830 of 267 municipalities, whereas nowadays it consists of only 77 municipalities.

²³ One may argue that unobserved natural advantages, which may attract people and firms, may be correlated over time. For example, natural resources may be at certain locations for centuries. Natural resources may be an important input in production for manufacturing firms, but this is certainly not the case for business services. We therefore are not too concerned that unobserved natural advantages of the present time are correlated with natural advantages in 1830.

5. Results

We present the main results employing parametric and semiparametric approaches to estimate the sorting model in Table 2, where we focus on the latter results. The parametric approach implies that coefficients are identical for different firms. First, we do not allow for endogeneity, so we estimate an ordinary Poisson model. We also estimate parametric models employing the control function approach. The four instruments have sufficient predictive power: the F -values for both specialisation and employment density exceed 1,300, so our instruments appear to be very strong (and also have the expected signs). For more results of the first step, we refer to Appendix C.

Table 2: Results for the presence of business services firms

	Parametric Regression						Semiparametric Regression			
	Poisson			Poisson with Control Function			Weighted Poisson		Weighted Poisson with Control Function	
	Coeff.	Std. Error		Coeff.	Std. Error		Mean Coeff.	Std. Dev.	Mean Coeff.	Std. Dev.
Specialisation	0.341	(0.010)	***	0.501	(0.115)	***	0.546	0.185	0.394	0.432
Diversity	-0.078	(0.013)	***	-0.101	(0.022)	***	-0.147	0.126	-0.164	0.139
Employment Density (log)	0.522	(0.015)	***	-0.047	(0.133)		0.814	0.288	0.411	0.624
Distance to Highway Ramp	-0.056	(0.008)	***	-0.022	(0.012)	*	0.105	0.342	-0.055	0.327
Distance to Station	0.081	(0.007)	***	-0.012	(0.014)		0.129	0.231	-0.019	0.344
Distance to Water	-0.185	(0.034)	***	-0.437	(0.095)	***	-0.282	0.383	-0.170	0.319
Distance to Open Space	0.108	(0.035)	***	0.522	(0.286)	*	-0.036	0.429	0.617	1.060
Area Size (log)	0.198	(0.007)	***	0.051	(0.024)	**	0.186	0.048	0.146	0.089
Control Function Specialisation				-0.108	(0.117)				0.205	0.465
Control Function Employment Density				0.738	(0.139)	***			0.611	0.826
Geographic Coordinates Included (5)	Yes			Yes			Yes		Yes	
Pseudo R^2	0.1285									
Number of Locations	21,465			21,465			21,465		21,465	
Number of Firms	9,170			9,170			9,170		9,170	

NOTE: The dependent variable is the number of firms in each postcode 6-digit area. The standard errors are between parentheses for the parametric specifications. In the parametric specification with the control function we present bootstrapped standard errors.

For the semiparametric approach, we relax the assumption on homogeneity (as explained in Section 3). In line with the literature, we set the window size to 10 percent of the total locations, so $k = 2,147$, to get reliable estimates of the marginal effects.²⁴ The average effects over all 9,170 firms of an increase for the

²⁴ The average distance of a business services firm to all other business services firms within the window is 5.07 kilometres. The average maximum distance is 9.04 kilometres. Pagan and Ullah (1999) and McMillen (2004; 2010) argue that the window size should not be too small when the goal is to estimate marginal effects, although this may lead to an underestimate of the heterogeneity in the estimated coefficients. In the sensitivity analysis, we will show that lower window sizes lead to unreliable estimates (see also Redfearn, 2009).

standardised measures of specialisation and diversity are presented, as well as the standard deviations of these effects.²⁵ We provide the results of specifications where we do not correct for endogeneity and we provide results of specifications where we use the control function approach to correct for endogeneity.

In Subsection 5.1 and 5.2 we discuss the findings of the first stage. Subsection 5.3 investigates the second stage results. Subsection 5.4 and 5.5 test the robustness of the results with respect to respectively internal instruments and other specifications.

5.1 First stage results: average effects for specialisation and diversity

Specialisation is positively related with location decisions of business services firms. Between-sector diversity has a much smaller effect and is negatively related to profits of business services. Controlling for endogeneity leads to a *higher* coefficient for specialisation (but to a substantially *lower* effect for the control variable employment density). In the semiparametric approach, we see that specialisation has on average a significantly lower coefficient and a higher standard deviation when we employ a control function approach. Again, the average effect of diversity is hardly influenced by the inclusion of the control function.

Conditional on employment density, one unit increase in specialisation in business services increases the probability that a firm locates there with on average 38 percent, in line with Head et al. (1995), Guimarães et al. (2000), Figueiredo et al. (2002), Rosenthal and Strange (2003) and Barrios et al. (2006). Some other studies that focus on manufacturing firms find small or no effects of specialisation (see Guimarães et al., 2004; Woodward et al., 2006). It may be expected that *manufacturing* firms do not rely on other manufacturing firms nearby, but more on specialised business services in vicinity, as they offer access to local market information, technology and skills (Helmsing, 2001; Jacobs et al., 2010). In our study, however, access to specialist services is fully captured by our specialisation measure, which indeed is positively valued by business services firms. Our results suggest that knowledge probably mainly flows between firms within

²⁵ One may also be interested in the average effects over all 21,465 locations instead of 9,170 firms. It appears that these are very similar to the effects presented here.

the same sector (MAR-externalities). Other externalities such as labour market pooling and input sharing may also be important reasons why business services locate near other business services.

Between-sector diversity is negatively related to location decisions, but its effect is smaller. One unit increase in diversity decreases the probability that a firm chooses that location with on average about 16 percent.²⁶ This implies that, *ceteris paribus*, it is profitable to business services firms to locate near a group of firms that belong to the same sector, which makes sense. For example, one expects lobbying firms to locate close to governments, maritime service firms close to port-related manufacturing, etc.

The control variables have in general plausible signs. Firms prefer to locate in dense employment areas in terms of employment because of proximity to customers and suppliers. Locations near highway ramps are generally favoured, but the effect of railway stations is somewhat unclear. Locations near open water attract business services (consistent with the observation that the Maas river waterfront in Rotterdam is dominated by office buildings). The coefficient of open space is positive, indicating that locations near open space are less likely to attract firms, probably due to zoning restrictions. Larger areas are hosting more business services, *ceteris paribus*: 10 percent increase in area size increases the probability with 1.3 percent that a firm locates there.

5.2 First stage results: exploring heterogeneity in profit parameters

An important motivation of the current study is to pay attention to heterogeneity in coefficients for specialisation and diversity. Figures 3 and 4 present the effect for a unit increase in respectively specialisation and diversity.

²⁶ We also regressed the number of establishments on the shares of manufacturing firms, wholesale and retail, public services and other firms, conditional on employment density. It confirms that business services benefit more from within-sector clustering than from between-sector clustering, confirmed by the negative coefficient of the diversity index. However, since we are not able to correct for endogeneity (we lack a sufficient number of valid instruments) these results are suggestive, at best.

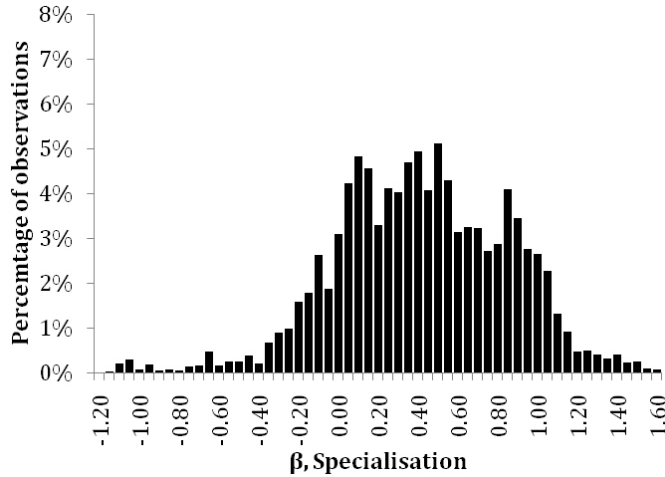


Figure 3: Distribution of effects for specialisation

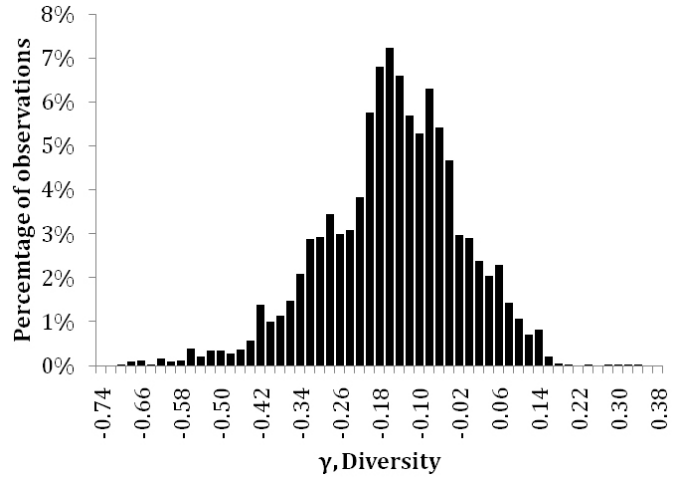


Figure 4: Distribution of effects for diversity

There is substantial heterogeneity in the effects for specialisation, 90 percent of the specialisation effects are between -0.40 and 1.00 and about 16 percent of the firms value specialisation negatively.²⁷ So, for most business services it is profitable to locate near other business services.²⁸ We also observe heterogeneity in the effects of diversity, although the spread is somewhat lower than for specialisation. 90 percent of the coefficients of diversity are between -0.32 and 0.14, 85 percent of the firms value diversity negatively.

It may also be worthwhile to look at the relation between the coefficients for specialisation and diversity. It appears that there is a strong negative correlation (of -0.49) between these coefficients (see Figure 5). Hence, our results imply that business services firms either rely on within-sector interactions or between-sector interactions: for relatively few firms profits depend negatively on both specialisation and diversity (about 11 percent) and for even fewer firms profits depend positively on both specialisation and diversity (about 6 percent).

²⁷ In the literature there is no consensus whether specialisation or diversity are important in firms' location decisions (see Duranton and Puga, 2000). One explanation is that previous studies do not take into account that the effects of specialisation and diversity on firm profitability are rather heterogeneous. As our results suggest that a substantial share of the firms value specialisation negatively, this may be important.

²⁸ We indeed see such local clustering: for example in the city centres of Rotterdam and The Hague, but also specialised clusters of only one type of business services have emerged. For example, maritime business services (lawyers, insurance companies) cluster in the Scheepvaartkwartier near the Maas river in Rotterdam (see Jacobs et al., 2010). These are often small firms that provide complementary knowledge-intensive services to firms located in the port.

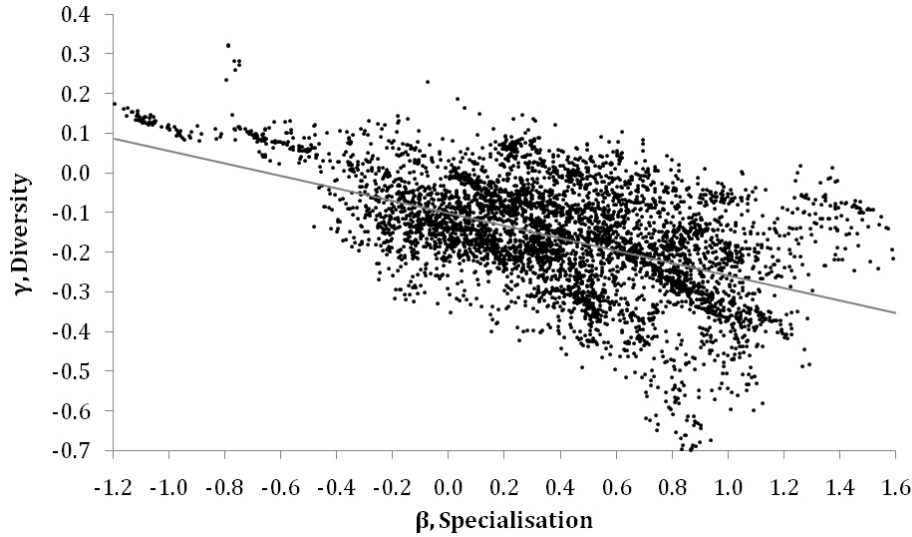


Figure 5: Relationship between effects for specialisation and diversity

5.3 Second stage results: explaining heterogeneity in profit parameters

In the previous stage, we have estimated the firm-specific effects of specialisation and diversity on location, which can be interpreted as the (normalised) effects on the profits of these firms. One may wonder whether or not one can explain the heterogeneity of these effects. In particular, we are interested in the effect of firm's workforce size, measured here by the number of employees. This is particularly important because we focus in our application on locations of *established* firms which are usually much larger than *new* firms that are the main focus of the current firm location choice literature. Duranton and Puga (2001) argue that new firms have to experiment with production processes and therefore prefer diversified locations. Once a firm finds its ideal production process, it relocates to a specialised location with firms that have the same production process.

So, we estimate (2) as explained in Section 3.3: the estimated firm-specific effects are regressed on firm characteristics, which consist of the logarithm of the number of employees as well as dummy indicators for

the six main types of business services (information technology services, financial services, legal services, general consultancy, real estate services and research and development).²⁹ Table 3 presents the results.

Table 3: Explaining heterogeneity in profit parameters by observable firm characteristics

	<i>Parametric Regression</i>							
	Specialisation		Diversity		Specialisation		Diversity	
	<i>Coeff.</i>	<i>Std. Error</i>	<i>Coeff.</i>	<i>Std. Error</i>	<i>Coeff.</i>	<i>Std. Error</i>	<i>Coeff.</i>	<i>Std. Error</i>
Workforce Size (log)	0.014	(0.004) ***	-0.004	(0.001) ***	0.010	(0.004) **	-0.003	(0.001) **
Age Index					-0.026	(0.005) ***	0.008	(0.001) ***
ICT Services	0.123	(0.038) ***	-0.021	(0.015)	0.130	(0.038) ***	-0.023	(0.015)
Financial Services	0.040	(0.038)	-0.028	(0.015) *	0.030	(0.038)	-0.025	(0.015) *
Real Estate Agencies	0.053	(0.039)	-0.041	(0.015) ***	0.047	(0.038)	-0.039	(0.015) ***
Legal Services	0.050	(0.038)	-0.034	(0.015) **	0.043	(0.037)	-0.032	(0.015) **
General Consultancy	0.061	(0.037) *	-0.030	(0.014) **	0.062	(0.036) *	-0.031	(0.014) **
Constant	0.300	(0.038) ***	-0.124	(0.015) ***	0.311	(0.038) ***	-0.128	(0.015) ***
R ²		0.0048		0.0028		0.0083		0.0061
Number of Firms		9,170		9,170		9,170		9,170

*NOTE: The dependent variables are the profit parameters of each firm for specialisation and diversity. The standard errors are between parentheses for the parametric specifications. The coefficients are significant at *0.10, **0.05 and ***0.01 levels.*

We find that larger business services firms benefit more from the presence of other business services firms (see Table 3). Our estimates imply that the effect of specialisation on profitability is about 15 percent higher for a large firm with 250 employees than for the average firm with 25 employees. This result is consistent with the firm sorting effect discussed by Baldwin and Okubo (2006) and more specifically in line with Duranton and Puga (2001) who showed that more mature firms tend to move towards specialised locations. Similarly, we find that it is more beneficial to larger business services firms to be closely located to groups of firms which belong to the same sector rather than close to a diverse group of firms (as we find that the effect of workforce size is negative on the diversity effect). The size of the workforce effect is similar to the effect on specialisation. This result seems also consistent with Henderson et al. (1995) and Duranton and Puga (2001), among others.

We also investigate the effects of firm age on profitability for specialisation and diversity more directly. We do not have detailed data on the age of the establishment, but each establishment has a unique ID and newer firms have a higher establishment ID number, which we use as an (imperfect) indicator for firm age.

²⁹ The dependent variable is not observed, but estimated. In this case, it is useful to rewrite the dependent variable as the sum of the value of the unobserved variable and measurement error (see Gawande, 1997). As long as this measurement error is random, the standard errors of the estimated effects in the second step are the correct ones.

We include this age index in the regression, transforming this index with zero mean and unit standard deviation. It is observed that this index is negatively and statistically significant related to the profitability of specialisation.³⁰ In line with theoretical predications Duranton and Puga (2001), we find that older firms have a stronger preference for specialised locations.

We furthermore find that there are some sector-specific differences in the profit parameters. For example, for ICT services the profitability of specialisation is significantly higher. For innovative R&D activities, we find that diversity is relatively important, which is consistent with Combes (2002) and Duranton and Puga (2001) who find that diversified locations attract innovative activities. So, overall, between-sector clusters are relatively more profitable for smaller innovative (R&D) firms.

5.4 Robustness: internal and external instruments

In the preceding sections we have used instruments that are external to the model. Timmins (2005) and Bayer and Timmins (2007) show that one may also use internal instruments, which naturally arise out of the sorting model. Timmins (2005) and Bayer and Timmins (2007) argue that firm's demand for a certain location j is not only dependent on its own attributes, but also on how it fits in the broader landscape. Therefore, exogenous attributes of other locations may act as proper instrument, as these variables influence the locational equilibrium but do not have a direct impact on profits (see also Klaiber and Phaneuf, 2010). Because of the number of exogenous attributes of other locations is large, Bayer and Timmins (2007) create a single instrument, which is equal to the predicted employment share of each location when coefficients related to the endogenous variables are set to zero.

Following this idea, we regress employment and business services employment (for locations where employment is located) on the average values of exogenous attributes between 2.5 and 3.5 kilometres of

³⁰ Although the R^2 is low in general, we see a relatively strong increase in the R^2 when we include the age index.

location j , using a Poisson model.³¹ Using the estimated coefficients, we estimate the predicted number of employees and employees in the business services sector for each location j and estimate the employment density $\sum_G \hat{\omega}_{Gj}$ and the business services specialisation measure $\hat{\rho}_j$, which are used as instruments. This approach relies on cross-sectoral variation in coefficients for locational attributes to ensure variation in the instruments, as emphasised by Bayer and Timmins (2007). Without any effective variation in geographic coefficients, the parameter estimates are not robust to misspecification of the profit function or the distributional assumptions of the error term. In our application, distributions of the coefficients to be estimated are nonparametric, so variation in coefficients is explicitly allowed for (and ensured as seen in Section 5.2).

Using the internal instruments, we find that the average effect of specialisation is 0.470 with a standard deviation of 0.474 and the average effect of diversity is -0.196 with a standard deviation of 0.192. Although the effect of specialisation is slightly higher (closer to the estimates without instruments), both the internal and external instruments approaches lead to the same conclusion. Figures 6 and 7 present the coefficients for both approaches. We observe reasonably high correlations between coefficients for specialisation (0.65)

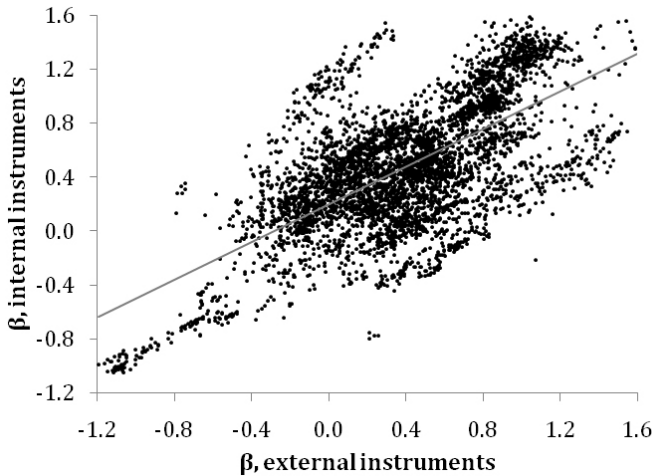


Figure 6: Coefficients for specialisation
with internal vs. external instruments

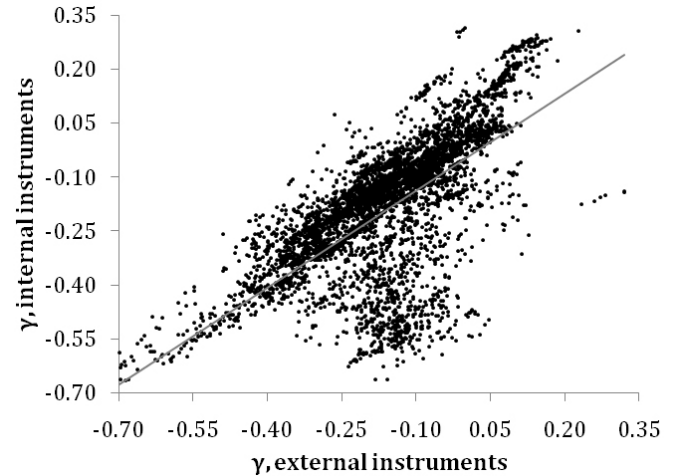


Figure 7: Coefficients for diversity
with internal vs. external instruments

³¹ So, the direct effect of employment that is located farther than 2.5 kilometres of the location is assumed to have a negligible effect on location choices. This is a reasonable assumption. For example, for $\delta = 2.5$, on average about 99 percent of the weight is within 2.5 kilometres of location j .

and between coefficients for diversity (0.64). So, our findings based on the external instruments specification are similar to the results with the internal instruments specification.

5.5 Robustness: other specifications

We also will demonstrate that our results are robust to changes in specifications and estimation procedures.

Table 4 summarises the results.

Table 4: Robustness checks for impact of specialisation and diversity on presence of business services firms

	GLM-IV			Distance Decay, $\delta=1$		Distance Decay, $\delta=5$		No Area Size		No Diversity	
	Coeff.	Std. Error		Mean Coeff.	Std. Dev.	Mean Coeff.	Std. Dev.	Mean Coeff.	Std. Dev.	Mean Coeff.	Std. Dev.
Specialisation	0.460	(0.103)	***	0.479	0.746	0.478	0.477	0.396	0.443	0.355	0.463
Diversity	-0.222	(0.023)	***	-0.392	0.467	-0.027	0.132	-0.240	0.164		
Number of Firms		9,170			9,170		9,170		9,170		9,170

NOTE: See Table 2.

First, we compare our parametric control function approach with a two-step instrumental variables Poisson model using the generalised linear model approach as proposed by Hardin et al. (2003) and Carroll et al. (1995). We see that the estimated coefficients are very comparable to those presented in Table 2, although the effect of diversity is somewhat more negative. Second, in this study we assumed that the distance decay parameter δ is equal to 2.5, so that most of the weight of the employment density measure is within 1.5 kilometres. We check whether the choice of δ influences our results. It appears that specialisation has a somewhat larger impact for both higher and lower values of δ , so our estimate may be somewhat conservative. For $\delta = 1$, the effect of diversity becomes almost twice as large, suggesting that the negative effect of diversity is stronger over longer distances. When we set $\delta = 5$, so that most of the weight of the employment is within a few hundred meters, the negative effect of diversity almost disappears. This suggests that potential benefits of diversity are primarily local or that costs of diversity are primarily regional.

Third, we check whether the results are influenced by excluding the potentially endogenous variable area size. We see that the effects are very similar. Excluding the diversity measure leads to almost the same result

for specialisation. We also verify whether our results are robust to the number of locations considered in the choice set of firms, so to the choice of window size. Figures 8 and 9 display the results.

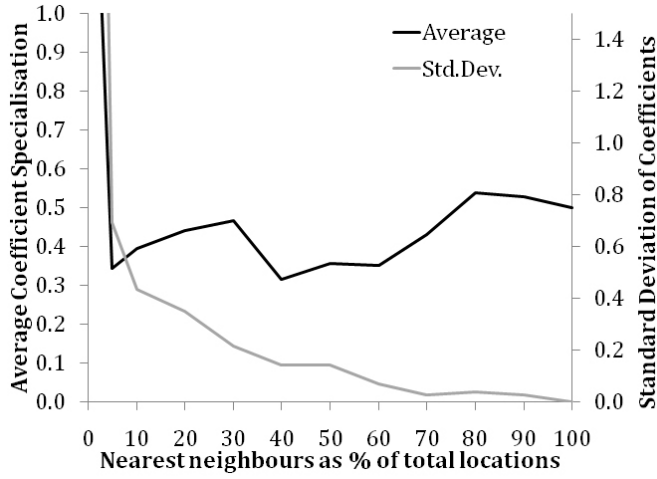


Figure 8: Average effect of specialisation the window size

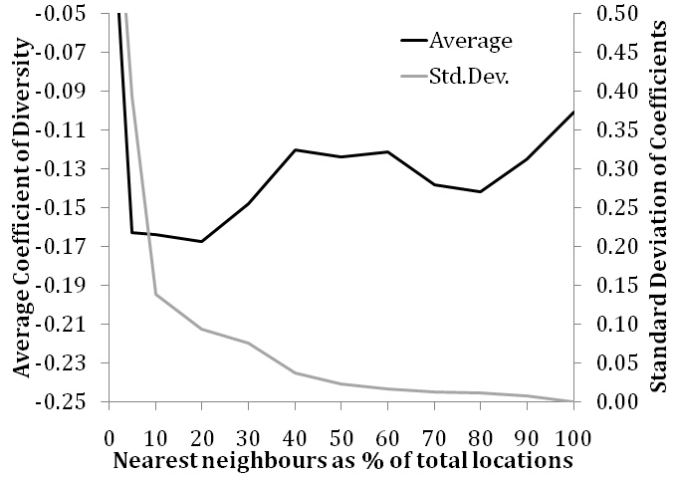


Figure 9: Average effect of diversity given the window size

Except when the number of locations in the choice set is very small, the average effects of specialisation and diversity are respectively between 32 and 54 percent and -10 percent and -17 percent. For a larger window size k , the effect of diversity is closer to zero. When the number of nearest neighbours k becomes too small, we cannot provide a reliable estimate of the effects of specialisation and diversity on firms' location decisions, because the standard deviations of the specialisation and diversity coefficients are too large. More specifically, Figures 8 and 9 reveal that for window sizes smaller than 2,147 observations (or 10 percent of the total number of locations), the standard deviation becomes unrealistically large.³²

6. Conclusion

In this paper we investigate the effects of local interactions, and in particular specialisation and diversity, on profitability, and therefore location decisions, of established business services firms in Zuid-Holland, a region in the west of the Netherlands. We pay special attention to firm-specific effects, as it is emphasised in the theoretical literature that firm-specific effects are relevant (Baldwin and Okubo, 2006). This argument is

³² When k is equal to the total number of locations, the estimated average effect is equal to the estimates of the parametric specification with the control function. The standard deviation is then, obviously, equal to zero.

particularly relevant for established rather than new firms that receive much attention in the literature. We estimate a heterogeneous sorting model for firms where firm-specific profit coefficients can be identified given information on firm location choice. We make a distinction between these local interactions effects and unobserved natural advantages and show that the results are robust with respect to the choice between external and internal instruments.

The results show that specialised clusters of business services have a strong positive effect on profitability of business services firms. A standard deviation increase in specialisation will, on average, lead to a 39 percent increase in the probability that a firm chooses that location. Between-sector diversity is, in general, negatively related to firm location decisions, so it is more profitable for business services firms to locate close to one type of sector (e.g. government), than to a range of different sectors. We also show that there is considerable heterogeneity in how specialisation and diversity affect location decisions: 84 percent of firms positively value locations with high levels of specialisation, whereas only 11 percent positively value diversified areas. It appears that there is a strong negative relationship between the effects of specialisation and diversity on profitability, so business services firms rely either on within-sector interactions or on between-sector interactions. Our results provide strong support for theories of Marshall, Arrow and Romer: local interactions mainly take place between firms of the same sector. Interactions between firms belonging to different sectors –Jacobs type externalities– are important for only a limited number of firms in the business services sector. More specifically, we show that smaller innovative firms gain more from diversity. Within-sector clustering is particularly profitable for large mature firms. Both these results are in line with theoretical predictions of Duranton and Puga (2001).

References

- Arrow, K.J. (1962). The Economic Implications of Learning by Doing. *Review of Economic Studies* 29: 155-173.
- Arzaghi, M., Henderson, J.V. (2008). Networking off Madison Avenue. *Review of Economic Studies* 75: 1011-1038.
- Audretsch, D.B., Feldman, M.P. (1996). R&D Spillovers and the Geography of Innovation and Production. *American Economic Review* 86(3): 630-640.
- Bajari, P., Benkard, C.M. (2005). Demand Estimation with Heterogeneous consumers and Unobserved Product Characteristics: A Hedonic Approach. *Journal of Political Economy* 113(6): 1239-1276.

- Bajari, P., Kahn, M.E., (2005). Estimating Housing Demand with an Application to Explaining Racial Segregation in Cities. *Journal of Business and Economic Statistics* 23(1): 20-35.
- Baptista, R., Escária, V., Madrugá, P. (2008). Entrepreneurship, Regional Development and Job Creation: the Case of Portugal. *Small Business Economics* 30: 49-58.
- Baldwin, R., Okubo, T. (2006). Heterogeneous Firms, Agglomeration and Economic Geography: Spatial Selection and Sorting. *Journal of Economic Geography* 6(3): 323-346.
- Barrios, S., Görg, H., Strobl, E. (2006). Multinationals' Location Choice, Agglomeration Economies, and Public Incentives. *International Regional Science Review* 29(1): 81-107.
- Bartik, T. (1985). Business Location Decisions in the United States: Estimates of the Effects of Unionization, Taxes, and Other Characteristics of States. *Journal of Business and Economic Statistics* 3(1): 1-14.
- Bayer, P., Ferreira, F., McMillan, R. (2007). A Unified Framework for Measuring Preferences for Schools and Neighbourhoods. *Journal of Political Economy* 115(4): 588-638.
- Bayer, P., Timmins, C. (2007). Estimating Equilibrium Models of Sorting Across Locations. *The Economic Journal* 117: 353-374.
- Black, S.E. (1999). Do Better Schools Matter? Parental Evaluation of Elementary Education. *Quarterly Journal of Economics* 114: 577-599.
- Blundell, R., Powell, J.L. (2003). Endogeneity in Nonparametric and Semiparametric Regression Models. In: Dewatripont, M., Hansen, L.P., Turnovsky, S.J. (eds). *Advances in Economics and Econometrics: Theory and Applications*. Cambridge: Cambridge University Press.
- Blundell, R., Powell, J.L. (2004). Endogeneity in Semiparametric Binary Response Models. *Review of Economic Studies* 71: 655-679.
- Carroll, R. J. , Stefanski, L. A. (1990). Approximate Quasi-Likelihood Estimation in Models with Surrogate Predictors. *Journal of the American Statistical Association* 85: 652-663.
- Cheshire, P. Sheppard, S. (1995). On the Price of Land and the Value of Amenities. *Economica* 62: 247-267.
- Cho, S.-H., Lambert, D.M., Kim, S.G., Jung, S.H. (2009). Extreme Coefficients in Geographically Weighted Regression and Their Effects on Mapping. *GIScience and Remote Sensing* 46(3): 273-288.
- Ciccone, A., Hall, R.E. (1996). Productivity and Density of Economic Activity. *American Economic Review* 86(1): 54-70.
- Combes, P.P. (2000). Economic Structure and Local Growth: France 1984-1993. *Journal of Urban Economics* 47(3): 329-355.
- Combes, P.P., Duranton, G., Gobillon, L. (2008). Spatial Wage Disparities: Sorting Matters! *Journal of Urban Economics* 63(2): 723-742.
- Defever, F. (2006). Functional Fragmentation and the Location of Multinational Firms in the Enlarged Europe. *Regional Science and Urban Economics* 36: 658-677.
- Devereux, M.P., Griffith, R., Simpson, H. (2007). Firm Location Decisions, Regional Grants and Agglomeration Externalities. *Journal of Public Economics* 91: 413-435.
- Drennan, M.P., Kelly, H.F. (2010). Measuring Urban Agglomeration Economies with Office Rents. *Journal of Economic Geography*: 1-27
- Duranton, G., Overman, H.G. (2005). Testing for Localization Using Micro-Geographic Data. *Review of Economic Studies* 72: 1077-1106.
- Duranton, G., Puga, D. (2000). Diversity and Specialisation in Cities: Why, Where and When Does it Matter? *Urban Studies* 37(3): 533-555.
- Duranton, G., Puga, D. (2001). Nursery Cities: Urban Diversity Process Innovation, and the Life Cycle of Products. *American Economic Review* 91: 1454-1477.
- Ekeland, I., Heckman, J.J., Nesheim, L. (2004). Identification and Estimation of Hedonic Models. *Journal of Political Economy* 112(1): S60-S109.

- Ellison, G., Glaeser, E.L. (1999). The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration? *American Economic Review* 89(2): 311-316.
- Ellison, G., Glaeser, E.L., Kerr, W.R. (2010). What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. *American Economic Review* 100: 1195-1213.
- Fan, J., Heckman, N.E., Wand, M.P. (1995). Local Polynomial Kernel Regression for Generalised Linear Models and Quasi-Likelihood Functions. *Journal of the American Statistical Association* 90(429): 141-150.
- Figueiredo, O., Guimarães, P., Woodward, D. (2002). Home-Field Advantage.: Location Decisions of Portuguese Entrepreneurs. *Journal of Urban Economics* 52(2): 341-361.
- Fotheringham, A.S., Brunsdon, C., Charlton, C. (2002). Geographically Weighted Regression: the Analysis of Spatially Varying Relationships. *Chichester: John Wiley & Sons*.
- Fujita, M., Ogawa, H. (1982). Multiple Equilibria and Structural Transition of Non-Monocentric Urban Configurations. *Regional Science and Urban Economics* 12: 161-196.
- Fujita, M., Krugman, P.R., Venables, A.J. (1999). The Spatial Economy: Cities, Regions and International Trade. *Cambridge, MA: MIT Press*.
- Gabe, T.M., Bell, K.P. (2004). Tradeoffs Between Local Taxes and Government Spending as Determinants of Business Locations. *Journal of Regional Science* 44(1): 21-41.
- Gawande, K. (1997). Generated Regressors in Linear and Nonlinear Models. *Economics Letters* 54(2): 119-126.
- Glaeser, E.L., Kahn, M.E., Rappaport, J. (2008). Why do Poor Live in the Cities: the Role of Public Transportation. *Journal of Urban Economics* 63(1): 1-24.
- Glaeser, E.L., Kallal, H.D., Scheinkman, J.A., Shleifer, A. (1992). Growth in Cities. *Journal of Political Economy* 100(6): 1126-1152
- Guevara, C.A., Ben-Akiva, M. (2006). Endogeneity in Residential Location Choice Models. *Transportation Research Record: Journal of the Transportation Research Board* 1977: 60-66.
- Guimarães, P., Figueiredo, O., Woodward, D. (2000). Agglomeration and the Location of Foreign Direct Investment in Portugal. *Journal of Urban Economics* 47: 115-135.
- Guimarães, P., Figueiredo, O., Woodward, D. (2003). A Tractable Approach to the Firm Location Decision Problem. *The Review of Economics and Statistics* 85(1): 201-204
- Guimarães, P., Figueiredo, O., Woodward, D. (2004). Industrial Location Modelling: Extending the Random Utility Framework. *Journal of Regional Science* 44(1): 1-20.
- Hardin, J.W., Schmiediche, H. Carroll, R.J. (2003). Instrumental Variables, Bootstrapping, and Generalised Linear Models. *The Stata Journal* 3(4): 351-360.
- Head, K., Ries, J., Swenson, D. (1995). Agglomeration Benefits and Location Choice: Evidence from Japanese Manufacturing Investments in the United States. *Journal of International Economics* 38: 223-247.
- Heckman, J.J. (1979). Sample Selection Bias as a Specification Error. *Econometrica* 47: 931-959.
- Henderson, J.V., Kuncoro, A., Turner, M. (1995). Industrial Development in Cities. *Journal of Political Economy* 103(5): 1067-1090.
- Helmsing, A.H.J. (2001). Externalities, Learning and Governance: New Perspectives on Local Development. *Development and Change* 32: 277-308.
- Holl, A. (2004). Manufacturing Locations and Impacts of Road Transport Infrastructure: Empirical Evidence from Spain. *Regional Science and Urban Economics* 34(3): 341-363.
- Ioannides, Y.M., Zabel, J.E. (2008). Interactions, Neighbourhood Selection and Housing Demand. *Journal of Urban Economics* 63: 229-252.
- Jacobs, J. (1969). The Economy of Cities. *New York: Vintage*.
- Jacobs, W., Koster, H.R.A., Hall, P.V. (2010). The Location and Global Network Structure of Maritime Advanced Producer Services. *GaWC Research Bulletin* 342, forthcoming in *Urban Studies*.
- Jaffe, A.B. (1989). Real Effects of Academic Research. *American Economic Review* 79: 957-970.

- Klaiber, H.A., Phaneuf, D.J. (2010). Valuing Open Space in a Residential Sorting Model of the Twin Cities. *Journal of Environmental Economics and Management* 60: 57-77.
- LeRoy, S.F., Sonstelie, J. (1983). Paradise Lost and Regained: Transportation Innovation, Income, and Residential Location. *Journal of Urban Economics* 13: 67-89.
- Lucas, R.E., Jr. (2001). Externalities and Cities. *Review of Economic Dynamics* 4: 245-274.
- Lucas, R.E., Jr., Rossi-Hansberg, E. (2002). On the Internal Structure of Cities. *Econometrica* 70(4): 1445-1476.
- Manski, C.F. (2000). Economic Analysis Of Social Interactions. *Journal of Economic Perspectives* 14(3): 115-136.
- Marshall, A. (1920). Principles of Economics. London: MacMillan.
- McMillen, D.P., McDonald, J.F. (1998). Suburban Employment and Employment Density in Metropolitan Chicago. *Journal of Urban Economics* 43: 157-180.
- McMillen, D.P. (2010). Issues in Spatial Data Analysis. *Journal of Regional Science* 50(1): 119-141.
- McMillen, D.P., Redfearn, C.L. (2010). Estimation and Hypothesis Testing for Nonparametric Hedonic House Price Functions. *Journal of Regional Science* 50(3): 712-733.
- McMillen, D.P. (2004). Employment Densities, Spatial Autocorrelation, and Subcentres in Large Metropolitan Areas. *Journal of Regional Science* 44(2): 225-243.
- Muller, E., Zenker, A. (2001). Business Services as Actors of Knowledge Transformation: the Role of KIBS in Regional and National Innovation Systems. *Research Policy* 30(9): 1501-1516.
- Pagan, A., Ullah, A. (1999). Nonparametric Econometrics. New York: Cambridge University Press.
- Petrin, A., Train, K. (2010). A Control Function Approach to Endogeneity in Consumer Choice Models. *Journal of Marketing Research* 47: 3-13.
- Porter, M.E. (1990). The Competitive Advantage of Nations. New York: Free Press.
- Redfearn, C.L. (2009). How Informative are Average Effects? Hedonic Regression and Amenity Capitalisation in Complex Urban Housing Markets. *Regional Science and Urban Economics* 39: 297-306.
- Rice, J. (1984). Bandwidth Choice for Nonparametric Regression. *Annals of Statistics* 12: 1215-1230.
- Romer, P.M. (1986). Increasing Returns and Long-Run Growth. *Journal of Political Economy* 94: 1002-1037.
- Rosenthal, S.S., Strange, W.C. (2003). Geography, Industrial Organisation, and Agglomeration. *Review of Economics and Statistics* 85(2): 377-393.
- Rosenthal, S.S., Strange, W.C. (2008). The Attenuation of Human Capital Spillovers. *Journal of Urban Economics* 64(2): 373-389.
- Schmidheiny, K., Brühlhart, M. (2011). On the Equivalence of Location Choice Models: Conditional Logit, Nested Logit and Poisson. *Journal of Urban Economics* 69: 214-222.
- Schumpeter, J. (1942). Capitalism, Socialism and Democracy. New York: Harper.
- Stute, W. (1984). Asymptotic Normality of Nearest Neighbour Regression Function Estimates. *Annals of Statistics* 12(3): 917-926.
- Timmins, C. (2005). Estimable Equilibrium Models of Locational Sorting and Their Role in Development Economics. *Journal of Economic Geography* 5: 83-100.
- Van Dijk, J., Pellenbarg, P.H. (2000). Firm Relocation Decisions in the Netherlands: An Ordered Logit Approach. *Papers in Regional Science* 79: 191-219.
- Van Stel, A., Suddle, K. (2008). The Impact of New Firm Formation on Regional Development in the Netherlands. *Small Business Economics* 30: 31-47.
- Woodward, D., Figueiredo, O., Guimarães, P. (2006). Beyond the Silicon Valley: University R&D and High-Technology Location. *Journal of Urban Economics* 2006: 15-32.
- Wheeler, D.C., Tiefelsdorf, M. (2005). Multicollinearity and Correlation among Local Regression Coefficients in Geographically Weighted Regression. *Journal of Geographic Systems* 7: 161-187.
- Yatchew, A. (2003). Semiparametric Regression for the Applied Econometrician. Cambridge: Cambridge University Press.

Appendix A. Equivalence relationship of locally weighted CLM and Poisson model

Assume that we have the same profit function as in (1):

$$(A1) \quad \max \pi_j^i = \alpha_0^i + V_j^i + \varepsilon_j^i,$$

where $V_j^i = \mathbf{X}_j' \boldsymbol{\alpha}^i + \beta^i \rho_j + \gamma^i \sigma_j$. Recall that we assumed that all firms at location j have the same V_j^i , so the local likelihood for each firm i at j may be defined as (see Fan et al., 1995):

$$(A2) \quad \log L^{i,logit} = \log L_j^{logit} = \sum_{j=1}^J w_{jj} n_j \log \frac{e^{\alpha_{0j} + V_j}}{\sum_{j=1}^J e^{\alpha_{0j} + V_j}} = \sum_{j=1}^J w_{jj} n_j \log \frac{e^{V_j}}{\sum_{j=1}^J e^{V_j}},$$

where w_{jj} denotes a local weight and n_j denotes the number of firms at a certain location j . In the local regression of j , α_{0j} is a constant. Now assume that $w_{jj} = \{0,1\}$, which encompasses a nearest neighbour approach; $w_{jj} = 1$ for the k nearest locations, where $k < J$ and zero otherwise. Then the likelihood becomes:

$$(A3) \quad \log L_j^{logit} = \sum_{j=1}^k n_j \log \frac{e^{V_j}}{\sum_{j=1}^k e^{V_j}}.$$

Maximisation of (A3) with respect to V_j shows that for $j > k$ the first-order condition $\partial \log L_j^{logit} / \partial V_j = 0$ is never fulfilled (as V_j , $j > k$ appears in the denominator of (A3)). So for $j > k$, $V_j \rightarrow -\infty$, and therefore $e^{V_j} = 0$. Hence, we may write (A3) as:

$$(A4) \quad \log L_j^{logit} = \sum_{j=1}^k n_j \log \frac{e^{V_j}}{\sum_{j=1}^k e^{V_j}}.$$

An alternative approach to arrive at the same result is to assume that the number of firms n_j is independently Poisson distributed, so:

$$(A5) \quad E(n_j) = \chi_j = e^{\alpha_{0j} + V_j}.$$

The log likelihood function may then be written as:

$$(A6) \quad \log L_j^{poisson} = \sum_{j=1}^J w_{jj} (n_j \log \chi_j - \chi_j - \log n_j!).$$

The first-order condition with respect to α_{0j} can be written as $\sum_{j=1}^J w_{jj} (n_j - \chi_j) = 0$. Solving for α_{0j} , we obtain:

$$(A7) \quad \alpha_{0j} = \log(\sum_{j=1}^J w_{jj} n_j) - \log(\sum_{j=1}^J w_{jj} e^{V_j}).$$

The concentrated likelihood function then yields:

$$(A8) \log L_j^{poisson} = \sum_{j=1}^J w_{jj} n_j \log \left(\frac{e^{v_j}}{\sum_{j=1}^J w_{jj} e^{v_j}} \right) + \sum_{j=1}^J w_{jj} n_j \log \left(\sum_{j=1}^J w_{jj} n_j \right) - \sum_{j=1}^J w_{jj} n_j - \sum_{j=1}^J w_{jj} \log n_j!,$$

We now again make the assumption that $w_{jj} = \{0,1\}$. It can be easily seen that this assumption is not only sufficient but also necessary to establish the equivalence relation between locally weighted CLM and locally weighted Poisson. So:

$$(A9) \log L_j^{poisson} = \sum_{j=1}^k n_j \log \left(\frac{e^{v_j}}{\sum_{j=1}^k e^{v_j}} \right) + C,$$

where $C = \sum_{j=1}^k n_j \log \left(\sum_{j=1}^k n_j \right) - \sum_{j=1}^k n_j - \sum_{j=1}^k \log n_j!$. The first term of (A9) is identical to the likelihood of the conditional Logit model (A4). As the log likelihood functions for the locally weighted CLM and the locally weighted Poisson model are identical up to a constant, maximum likelihood yields identical parameter estimates, given the assumption that weights are either zero or one.

Appendix B. Descriptive statistics

Table B1: Descriptives

	Mean	Std.Dev.	Min	Max
<i>Dependent variables</i>				
Number of Firms	0.427	1.239	0.000	40
Number of Employees	11.204	81.766	0.000	4,194
<i>Independent variables</i>				
Specialisation (untransformed)	0.153	0.091	0.000	0.785
Specialisation	0.000	1.000	-1.733	7.253
Diversity (untransformed)	1.972	0.706	0.522	7.266
Diversity	0.000	1.000	-2.140	7.344
Employment Density	6,373.672	8,226.440	9.264	64,382.340
Distance to Ramp (in km)	2.951	2.240	0.027	15.686
Distance to Station (in km)	3.225	3.562	0.009	29.065
Distance to Water (in km)	0.352	0.276	0.000	2.409
Distance to Open Space (in km)	0.197	0.188	0.000	1.792
Area Size in ha	8.131	61.666	0.001	2175.337
<i>Instruments</i>				
Population Density per km ² 1830	965.680	2,567.992	7.337	19,934.830
Location Bombed in 1940	0.030	0.169	0.000	1.000
Distance to Station 1850 (in km)	8.737	7.770	0.015	42.274
Location Below Water 1830	0.016	0.126	0.000	1.000
Number of Locations		21,466		
Number of Firms		9,170		

NOTE: The employment density variable is the weighted number of jobs given $\delta = 2.5$, see equation (4)

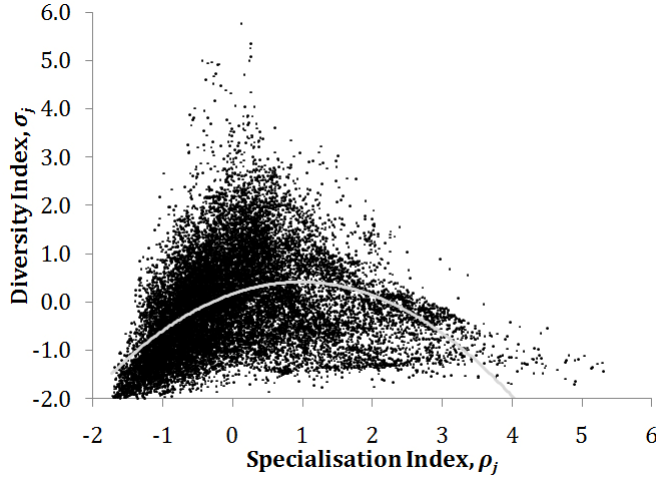


Figure B1: Relation between the specialisation measure and the diversity index

Appendix C. First stage estimation procedure

In this appendix we outline our estimation procedure in more detail. We also present some first-step results. The estimation procedure consists of the following steps. First, select appropriate instruments, external to the model, or internal to the model as proposed by Bayer and Timmins (2007). We use as external instruments area bombed in 1940, whether locations were below water in 1830, distance to stations in 1850 and the log population density 1830. Then, we regress specialisation and log employment densities on exogenous attributes of location and instruments, using the k -nearest neighbour approach. We set the window size to 10 percent. In Table C1 we present the first step results for the parametric and semiparametric specification. We see that we have very strong instruments: the F -values exceed 1,300. The instruments also have the expected signs. The effect of bombing is substantially larger in the parametric model, because in the semiparametric regression, many locations are not close to Rotterdam, and therefore the coefficient is equal to zero. The average coefficient is therefore also lower.

Table C1: First-step results

	Parametric Regression				Semiparametric Regression			
	Specialisation		Employment Density (log)		Specialisation		Employment Density (log)	
	<i>Coeff.</i>	<i>Std. Error</i>	<i>Coeff.</i>	<i>Std. Error</i>	<i>Mean Coeff.</i>	<i>Std. Dev.</i>	<i>Mean Coeff.</i>	<i>Std. Dev.</i>
Location Bombed in 1940	1.555	(0.039) ***	0.572	(0.034) ***	0.217	0.374	0.098	0.176
Location Below Water 1830	-0.110	(0.046) **	-0.123	(0.041) ***	-0.081	0.235	-0.117	0.268
Distance to Station 1850	0.000	(0.002)	-0.025	(0.001) ***	-0.105	0.447	-0.184	0.297
Population Density 1830 (log)	0.176	(0.004) ***	0.213	(0.004) ***	0.073	0.092	0.074	0.060
Other Variables Included (11)	Yes		Yes		Yes		Yes	
R^2	0.358		0.645					
F-Test for Instruments	1,347		1,387					
Number of Locations	21,465		21,465		21,465		21,465	
Number of Firms	9,170		9,170		9,170		9,170	

NOTE: See Table 2.

Third, using the first step estimates, we calculate for each location the residual of the regression of specialisation and employment density. We then include the residuals as control functions in the geographically weighted Poisson regression of number of business services firms per location on locational attributes, specialisation, diversity and employment density (see Petrin and Train, 2010). Using this approach, we correct for the endogeneity of specialisation and employment density.